Theoretical Framework

The following table lays out the links we have established between aspects of dialogic teaching and features of a dialogue system which could assist in the implementation of this approach. The right-hand column contains the features that we consider a dialogue system would need in dialogic teaching. The left-hand column contains the aspects of dialogic teaching most linked to those features. Our classification of the aspects of dialogic teaching was built around the principles described by Alexander (2010), which we link to the visible indicators described by Sedova (2017) and the oracy sub-skills detailed by Mercer (2017).

DIALOGIC TEACHING CONCEPTS	SYSTEM FEATURES
PRINCIPLE **Cumulative** **INDICATOR** **Uptake** **ORACY SKILL** • **Cognitive** • **Social and emotional** • **Linguistic**	Memory to store the main points of the conversation so far and build upon them; it will need to remember which arguments have already been presented and avoid repetition. Displaying chat logs so that the students can also remember what has been mentioned (Major et al, 2018; Mercer et al, 2010; Mercer, 2012).
	Generating questions on the topic of the discussion (Alexander, 2010), so it will need to identify the topic and have or gather information about it to make coherent contributions (Goda, 2014).
	Stance detection to analyze the students' arguments and present counterarguments and respond coherently (Goda, 2014; Stab, 2014; Kulatska, 2019).
	Encouraging students to use their own words to ensure that they understand the topic (Wegerif, 2020), but at the same time introducing appropriate terminology (Mercer, 2012).
PRINCIPLE **Collective** **INDICATOR** **Open discussion** **ORACY SKILL** **Social and emotional**	Measuring the length of the exchanges: neither the system nor the student can monopolize the conversation. Most exchanges should be somewhat lengthy to have substance (i.e. reasoning, building up on previous ideas…) (Sedova, 2017; Wegerif, 2020).
	Presenting a wide variety of ideas on the topic, for them to be evaluated and accepted or challenged; a student-dialogue system interaction could have a very limited perspective, which is why an effort should be made for the system to present more than one voice or point of view (Alexander, 2010; Wegerif, 2020).
	Matching the topic of the system's and the student's utterances: the speakers should not lead parallel conversations (Goda, 2014; Mercer et al, 2017; Wegerif, 2020).

PRINCIPLE **Purposeful**	Setting the goal of the exchange: it should not be mere chit-chat but reach some kind of understanding on a topic (Alexander, 2010; Sedova, 2017).
INDICATOR **Student reasoning**	Providing background information that can serve as the basis for the student's contribution (Kulatska, 2019).
ORACY SKILL **Cognitive**	Showing the student what can be achieved through the discussion; the task can be more successful if the students see value in it (Mercer, 2012; Thies, 2017).
	Feedback and access to chat logs to help the student reflect on their progress, also to increase accountability and motivation. (Mercer, 2010; Major, 2018; Huang, 2019).
	Making feedback salient and constant to increase motivation. Correcting the students on something that they did well but the system misinterpreted would cause frustration; however, if the students are not provided feedback throughout the task, they may not engage with it as desired. A good compromise might be to give frequent feedback which acknowledges that the student's work is in progress, that in open discussions there is not absolute right or wrong, and that the dialogue system is not infallible (Pinkwart, 2008; Kulatska, 2019).
	Identifying reasoning in the student's answers; it could be expressed implicitly (Jurafsky, 2019). When no reasoning is detected, the system should encourage reasoning (Mercer, 2010; Mercer, 2012; Sedova, 2017). The system should also model productive dialogue by reasoning (Mercer, 2010; Mercer, 2012; Okada, 2018).
PRINCIPLE **Reciprocal**	Countering students' arguments or reacting with further questions to delve deeper into the topic and develop ideas (Alexander, 2010; Wegerif, 2020).
INDICATOR • **Student questions** • **Open discussion** • **High-order questions**	Showing openness to questions by explicitly saying that it's open to questions, presenting information and telling the student to ask about it or presenting something new/strange/opposed to the student's view that may inevitably lead to questions (Alexander, 2010; Sedova, 2017).
ORACY SKILL • **Social and emotional** • **Cognitive**	Asking for clarification when the student's utterances are not understood (Jokinen, 2009; Bii, 2013). Giving explanations considering the level of understanding that the students have shown (Mercer, 2012).
	Dialogue systems may rely on the user taking charge of the conversation (user-initiative systems) or instead have it completely controlled by the system (system-initiative systems). A more dialogic exchange might require the more complex and human-like mixed-

	initiative design, where either participant (student or system) may direct the conversation at different points (Jurafsky, 2019).
	Helping the students see how much the dialogue system can understand. Users often address machines as they would people, so they need to be made aware that the dialogue system may understand less than a person, but also that they do not need to oversimplify their utterances. In the context of dialogic teaching, it might be especially important to show users that the dialogue system can process complex questions that go beyond fact regurgitation, so that students are encouraged to include those more cognitively demanding questions (Jokinen, 2009; Thies, 2017; Huang, 2019).
	Asking "high order questions", more linked to reflection than to information retrieval (Mercer et al, 2010b; Sedova, 2017). To discuss topics where many perspectives can be considered and questions have no single answer, the system would need very large and diverse data sources and a deep-learning approach to generate answers that seem natural and contribute to the discussion (Kuyven, 2018; Kulatska, 2019; Jurafsky, 2019).
PRINCIPLE **Supportive** **INDICATOR** **[Not directly observable]**	Language that, even when used to challenge students' ideas, is respectful and encouraging (Alexander, 2010; Mercer, 2010; Mercer et al, 2010b; Thies, 2017; Major, 2018; Ruan, 2019). A friendly avatar could be helpful (Bii, 2013; Huang, 2019), if its design is not distracting (Heller, 2016).
ORACY SKILL	Detecting confrontational talk and demanding that students express their ideas in a respectful, reasoned manner (Mercer et al, 2010).
• **Social and emotional** • **Linguistic** • **Physical**	Multimodality, to reinforces physical and linguistic skills and to adapt to different learning styles and accessibility needs (Jokinen, 2009; Heller, 2016; Major et al, 2018). Multimodality involves voice recognition and speech generation but may also involve a static or dynamic avatar (Bii, 2013; Heller, 2016).